International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Estimation of the Mass Density of Guayule from Height Data

Anderson Jr., R.J.

IIASA Working Paper

WP-81-070

June 1981

ESTIMATION OF THE MASS DENSITY OF GUAYULE FROM HEIGHT DATA

Robert J. Anderson, Jr.

June 1981
WP-81-70

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

PREFACE

In 1980, IIASA joined with the Centro de Investigacion en Quimica Aplicada (CIQA) to study resource development alternatives for arid and semi-arid regions. This joint effort is motivated by the perception that planning and programming of development projects, as they typically are applied to projects for drylands, are inadequate and pose serious obstacles to successful development of these regions.

Two characteristics distinguish the problem of planning and programming development projects for drylands. First, all of the common difficulties that beset development planning and programming (e.g., inadequate data, poorly understood social and cultural relations, inadequate infrastructure, inadequate organization capacity) are present in the extreme. Second, even modest-sized development projects are usually enormous in relation to the social, economic, and technical structure of drylands regions; their ramifications are little short of revolutionary.

To focus our efforts to improve planning and programming methods for dryland regions, it was decided to examine a specific problem: the prospects for developing a region in northern Mexico based on the exploitation of 6 vegetal resources native to the region. A description of this study is available in

> Anderson, R.J., E. Campos-Lopez, and D. Gourmelon. An Analysis of Renewable Resource Development Alternatives for the Northern Arid Region of Mexico: Study Prospectus. WP 81-7. International Institute for Applied Systems Analysis (January, 1981).

Guayule (parthenium argentatum gray) is one of the vegetal resources under investigation in the study. Guayule shrub, which grows wild on the sierras of the Chihuahuan Desert, produces a high molecular weight hydrocarbon that can be processed into a premium-quality rubber. For approximately 50 years during the first half of this century a small but important guayule rubber industry operated in Mexico.

The Mexican government currently plans to reactivate the guayule rubber industry using shrub harvested from wildstands as the basic source of shrub. This paper examines a critical aspect of this plan, the inventory of guayule shrub available for harvest. More specifically, the paper examines a new shrub inventory technique that has been proposed, and attempts to account for differences between the results obtained when this technique is used in place of standard inventory methods.

ACKNOWLEDGEMENTS

Helpful comments and suggestions from J. Ausubel, E. Campos-Lopez, J. Garcia-Delgado, A. John, J. Kindler, and E. Neavez-Camacho are gratefully acknowledged. The author is solely responsible for any errors, ambiguities, or other faults that may remain.

CONTENTS

ESTIMATION OF THE MASS DENSITY
OF GUAYULE FROM HEIGHT DATA

Robert J. Anderson, Jr.

1 INTRODUCTION

In their report entitled "Ecology and Reproduction of Guayule in Natural Stands in Coahuila, Nuevo Leon, and Zacatecas, Mexico," Native Plants, Incorporated [hereafter, NPI (1980)] estimates the mass density of guayule [i.e., the mass of guayule shrub per unit of land area] using an innovative short-cut technique based on an empirically determined relationship between plant height and weight, and data on the distribution of plant height and frequency of occurrence. Although it is impossible to make precise comparisons, the estimates of density obtained using NPI's method appear to be somewhat lower than have been reported in other studies.

According to NPI (1980), the stands in their sample with the highest percentage guayule canopy cover [i.e., between 12.1 and 25 percent coverage] had an estimated shrub density of 1.5 metric tons per hectare. The medium canopy cover stands [i.e., stands with canopy cover of between 5.8 and 12.0 percent] had an estimated shrub density of 1.01 tons of guayule shrub per hectare; and the low cover stands [i.e., between 4.2 and 5.6 percent] had an estimated 0.68 tons of shrub per hectare. These estimates contrast sharply with estimates provided in CONAZA

(1977) which imply average densities of about 1.64 metric tons per hectare over all stands in roughly the same region as was covered in NPI (1980). They contrast even more sharply with estimates implied by data presented in NPI (1981), which reports average densities on the order of 3.86 tons per hectare in an overlapping [but not identical] region. In general, the NPI (1980) estimate is lower than estimates reported in other sources.

The differences between these estimates are quite large [although reported data do not allow determination of whether the differences are statistically significant]. If we take the NPI(1980) estimate for medium cover stands as representative of an overall average density, then the NPI estimate [i.e., of 1.01 tons per hectare] is only about 62 percent as large as the CONAZA (1978) estimate [i.e., of 1.64 tons per hectare], and about 26 percent as large as the NPI (1981) estimate [i.e., of 3.86 tons per hectare].

Which estimate is most nearly correct? This is an important question. The economic viability of harvest of guayule from wildstands will depend strongly on the amount and spatial distribution of guayule shrub. Resolution of discrepancies in existing data on the amount and distribution of guayule is thus of the very highest priority.

This paper was motivated by the desire to explain discrepancies between the shrub density estimates described above. In a previous paper on the maximum sustainable yield from harvest of guayule wildstands [see Anderson (1981)], an effort was made to reconcile the findings of several inventories of guayule shrub. Some inventories found relatively small quantities of shrub [e.g., Lloyd (1911)], while others [e.g., CONAZA (1977) and companion inventories discussed in Anderson (1981)] found relatively large quantities. As a part of the reconciliation process, estimates of average shrub density were compared. As noted above, the estimates presented in NPI (1980) are somewhat lower than are found in other sources. It seemed possible when this investigation of density estimates began that understanding why the NPI (1980) estimates are lower than others might help to explain some of the differences between high and low estimates of shrub stock.

In attempting to reconcile density estimates, the NPI (1981) estimate was not considered since, according to the authors of that report, the procedure used resulted in an upward bias in estimated average shrub weight. The present paper does not explore all possible sources of the discrepancies between the NPI (1980) estimates and others. The information required to do so is not currently available. Instead, the paper concentrates on a potentially major source of bias: the method used in NPI (1980) to estimate shrub mass from shrub height data.

A subsidiary motivation in undertaking the analysis reported here was to examine the possibility of using short-cut methods to reduce the burden of conducting shrub inventories. A short-cut survey method like that proposed in NPI (1980) [this method will be described in detail in Section 2], if its validity could be demonstrated, would be a useful tool in reducing the amount of time and cost required to inventory the standing stock of guayule shrub in a region. The usual method of inventorying shrub is to select a large number of plots and to measure, pull, and weigh all adult individuals in each. This is, as noted above, both time consuming and costly. If an alternative could be found that reduced the number of shrubs that had to be measured, pulled, and weighed, without biasing or materially reducing the accuracy of the estimate, the burden of conducting shrub inventories would be reduced substantially.

The analysis presented here will show that the procedure used by NPI (1980) results in a systematic underestimate of shrub weight. Other things being equal, NPI's method always results in an estimated shrub weight that is lower than the true weight, and hence, an estimated density that tends to understate the true density. Moreover, it will be shown that this systematic bias could be large enough to account for the differences between the CONAZA (1977) and NPI (1980) estimates.

It must be stressed that the analysis in this paper uncovers a possible explanation for differences in these estimates. There could well be other explanations, and these too merit investigation. The most important conclusion to be drawn is that the methods and results of investigations bearing on the amount and distribution of guayule should be examined in depth and reconciled without delay.

While it cannot be concluded that the analysis in this paper definitively resolves differences between the NPI (1980)and CONAZA (1977) estimates, there is one important conclusion that is valid in all events: if a short-cut inventory method relating height and weight like that recommended in NPI (1980, p 31) is to be used, one must be careful to do the calculations in a way that does not impart systematic bias to the results. Calculation methods that do not impart such biases are developed below. These methods are likely to be quite useful inasmuch as the need for short-cut inventory techniques will grow with the start-up of commercial operations.

The plan of the paper is as follows. In Section 2, the procedure followed in NPI (1980) is reviewed step-by-step. This provides the background needed to understand why the NPI procedure results in a systematic underestimate of shrub mass.

Section 3 shows why the NPI procedure is biased and how the bias can be corrected. It also provides a simple-to-use approximating method that should be quite satisfactory in most applications and somewhat easier to compute than an exact calculation.

Section 4 uses data from NPI (1980) to estimate the possible size of the the bias in NPI's reported shrub density estimates. It will be shown that the possible bias could be large enough to account for the difference between the CONAZA (1977) and NPI (1980) estimates.

2 THE NPI PROCEDURE

The procedure followed in NPI (1980) involves the use of an empirically estimated relationship between guayule shrub height and shrub weight. The data for this empirical relationship were taken from a sample of 80 individual plants for which data were collected on weight, height, stem diameter, and crown diameter. The statistical relationship shown in Figure 1 was obtained by ordinary regression methods applied to the data collected on plant height and weight. Summary statistics reported in NPI (1980) are inadequate to appraise fully the extent to which the equation reported in Figure 1 is an adequate characterization of the sample data. Overall, the equation explains a little more than 68 percent of the variation in individual plant weight [i.e., $0.68 = 0.83^2$], and the relationship between height and weight is statistically significant at better than the 0.01 level.

Using this equation relating height and weight, and the average height of adult guayule plants [i.e., plants taller than 25 centimeters averaged over quadrats in which such plants occurred] in each plot surveyed, an estimate of the average weight of adult guayule was obtained. Estimated average weight was then multiplied by the observed frequency of guayule adult occurrence [obtained by dividing the number of adult plants by the total number of quadrats per plot] to obtain an estimated average mass per square meter [expressed in terms of kilograms per m^2]. Multiplication of this figure by 10 then resulted in an estimated mass per hectare, expressed as tons per hectare [i.e., kilograms per square meter times ten thousand square meters per hectare divided by 1000 kilograms per ton].

This procedure is easily clarified by means of an example.

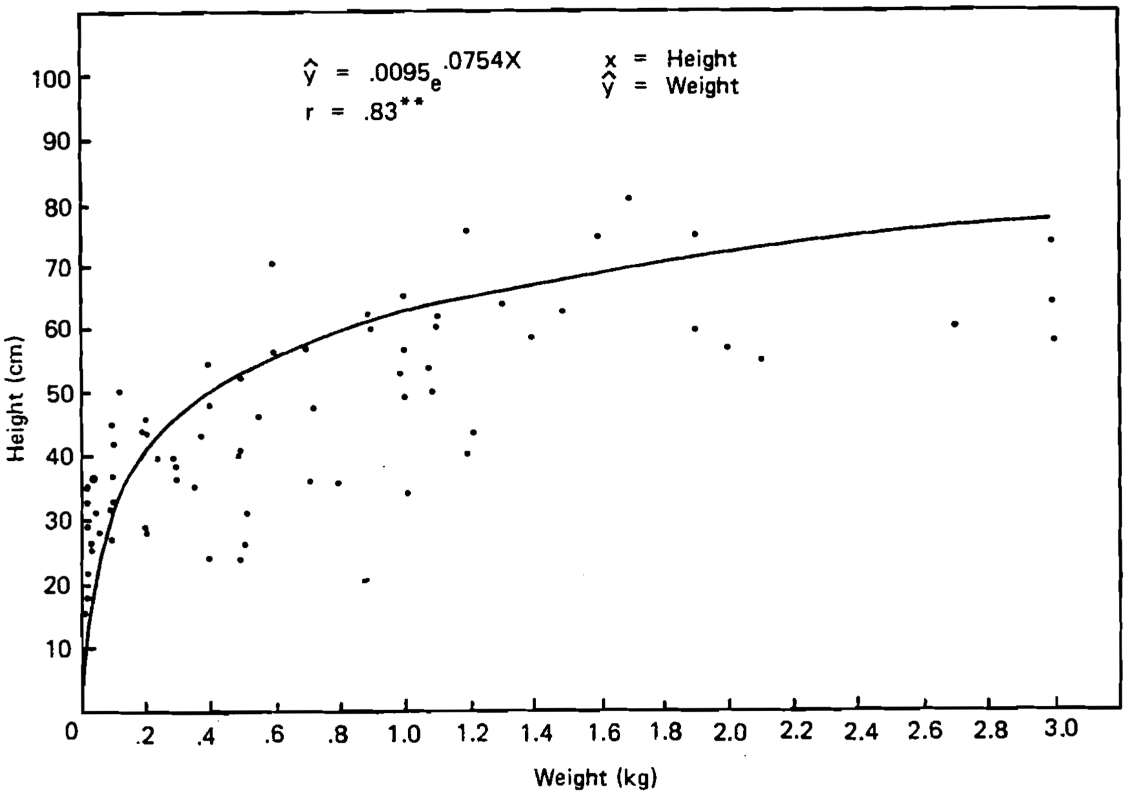

Figure 1. Relationship between height and weight of guayule individuals sampled. (Source: Figure 5, NPI, 1980. Reported coefficients corrected.)

**Significant at the one percent level.

Consider a plot in which the average height of guayule individuals is 40 centimeters and in which the observed frequency of occurrence of guayule 0.80. Substituting an average height of 40 centimeters into the equation reported in Figure 1, we obtain an estimated average shrub weight of 0.1939 kilograms. [i.e., $0.1939 = 0.0095e^{0.0754\ (40)}$]. Applying the assumed frequency of occurrence of 0.8, we obtain an overall estimate of density of 0.1551 kilograms per square meter, or 1.551 tons per hectare. This is precisely the calculation that is behind the estimated shrub densities reported in NPI (1980).

The NPI procedure is both simple and straightforward. It also holds forth the promise of materially reducing the time and money cost of conducting an inventory of shrub. NPI (1980, pp 27-31) recommends its usage for calculation of inventories according to the following [paraphrased] protocol:

> A 200 meter transect is walked and the following information is recorded: (1) Number of times that guayule occurs within a one square meter quadrat centered on one's right foot, denoted by N; (2) Average height of guayule adults in quadrats of occurrence, denoted by H. Then the formula illustrated above is applied to these data, using as an estimate of frequency N divided by 100.

In spite of its appealing simplicity and the potential it offers for reducing the time and money costs of shrub inventories, the NPI procedure leads to estimates of shrub density that are systematically in error. Even given absolutely accurate data, it leads to estimated shrub weights and densities that are lower than the true weights and densities.

3 BIAS IN THE NPI PROCEDURE

To demonstrate the bias in the NPI procedure, assume that the height-weight relationship shown in Figure 1 is absolutely accurate [i.e., that all guayule individuals obey exactly this relationship]. It will be shown that the bias in the NPI procedure is a result of improper use of this relationship.

To begin, the curve shown in Figure 1 has been redrawn with the axes switched, so that height now appears on the horizontal axis and weight appears on the vertical axis. This is done in Figure 2. Following this convention will make it easier to visualize each step in the analysis.

The bias in NPI's procedure can most easily be shown if it is assumed, without loss of generality, that all guayule plants are of one of two heights, "tall" and "short". Suppose that tall plants are 60 centimeters high and that short plants are 20 centimeters high. [Ignore, for purposes of this example, the fact that 20 centimeters is beneath the cutoff height for adult plants]. Further, suppose that exactly one half of all guayule plants are tall, and one half are short. The average height of plants would thus be 40 centimeters as in the example considered in Section 2. Recall from the calculations in Section 2 that the estimated average shrub weight corresponding to this average height, using the NPI procedure, is 0.1939 kilograms.

Instead of computing average height and substituting average height in the height-weight equation in Figure 1 [as was done in the calculations in Section 2], let us examine what

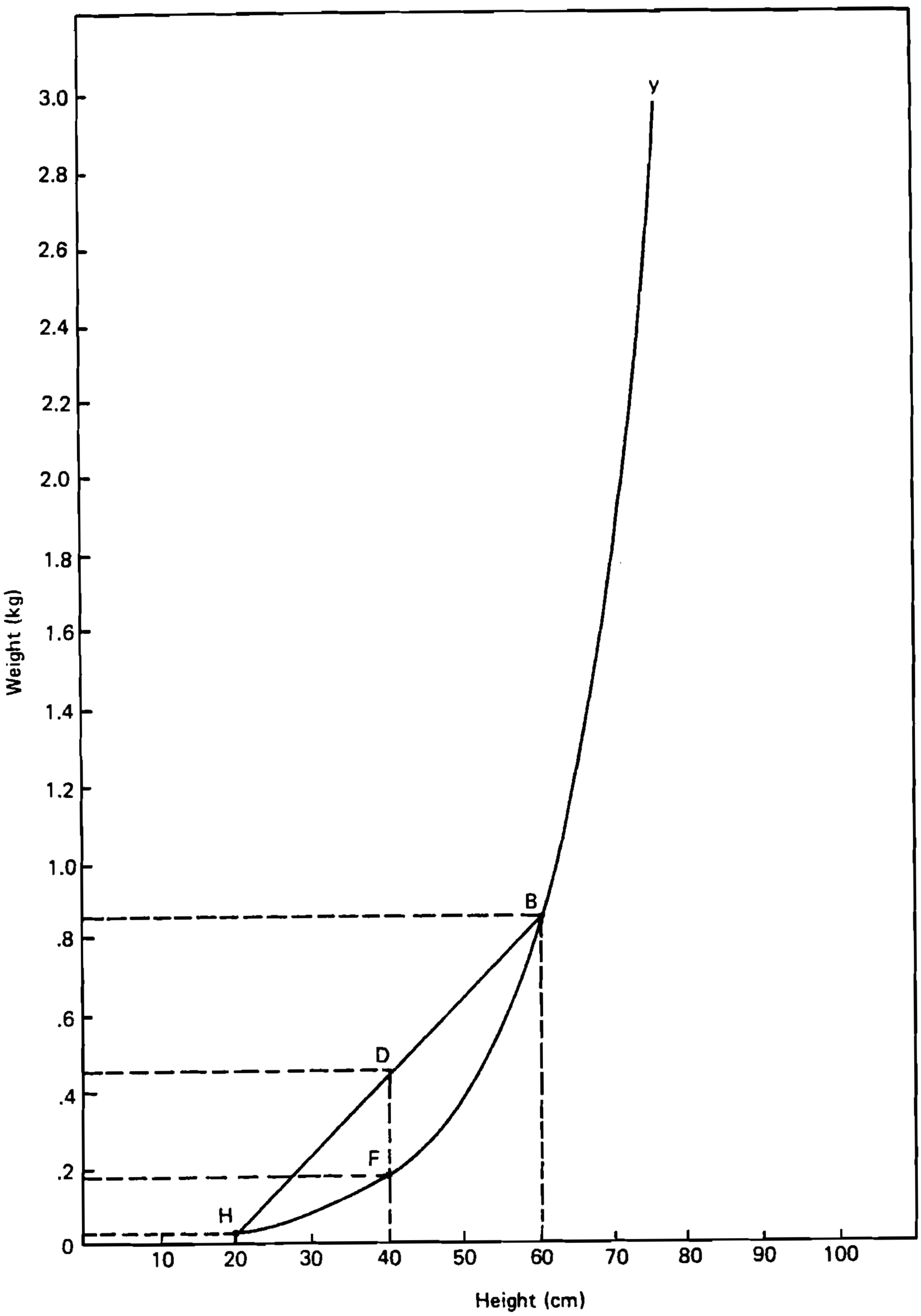

Figure 2. Example analysis.

happens if the heights of each size of plant are first substituted into the height-weight equation, thus obtaining an estimated weight for each size, and these estimated weights are then averaged. The weight of tall plants would be 0.8759 kilograms [i.e., $0.8759= 0.0095e^{0.0754 (60)}$] and the weight of short plants would be 0.0429 kilograms [i.e., $0.0429=0.0095e^{0.0754 (20)}$]. Examining Figure 2, these are the points B and H that would be read off of the curve 0-Y vertically above the points corresponding to 60 centimeters and 20 centimeters respectively.

By assumption, half of the guayule plants are 60 centimeters high, and half are 20 centimeters high. This means that half weigh exactly 0.8759 kilograms, and half weigh exactly 0.0429 kilograms [recall that we are assuming that the weights calculated using the height-weight relationship are exact]. The true average weight of a shrub would thus be 0.4594 kilograms [i.e., 0.4594=0.5(0.0429)+0.5(0.8759)].

This is clearly a much larger estimate of average weight than was obtained by computing mean plant height and substituting in the height-weight equation [recall that an estimate of 0.1939 kilogram was obtained by this procedure]. The reason that the estimates differ can also easily be seen by reference to Figure 2. The estimate corresponding to the procedure recommended by NPI is obtained by finding the point on the 0-Y curve that corresponds to the computed average height of 40 centimeters. This point is represented by point F in Figure 2. The estimate obtained using the procedure introduced in this section corresponds to finding the points on the 0-Y curve associated with the two plant sizes [shown by points B and H in the figure] and averaging the computed weights thus calculated, obtaining an average that is represented by point D in Figure 2.

The fact that the two calculations considered in this paper [i.e., the procedure recommended by NPI and the alternative calculation explained above] give different results is no mere matter of chance dependent upon the example chosen. Whenever weight is a convex function of height [i.e., a function whose slope increases with increasing values of height, as does the function depicted in Figure 2], use of the NPI procedure will result in a systematic underestimate of shrub weight and density. This can be proven easily by appeal to a fundamental result in mathematical statistics known as Jensen's Inequality [see Rao (1965) for a discussion of this result].

The most direct procedure to use to correct the error in NPI's procedure is to reverse the order of the calculations, as has been done in computing the example shown in Figure 2. That is, using data on individual plant heights, weights of individual plants should be computed first and then averaged, instead of averaging height first and then computing weight.

In cases in which this is impractical either due to the burden of computation or to the unavailability of the basic data on individual plant height, it may be possible to use an approximation based on summary statistics of plant height. This approximation will yield a more nearly correct answer than the NPI procedure. The approximation relies on the assumption that the height-weight function may be approximated adequately by a second-order Taylor series. Letting W(h) denote the height-weight function, this function may be approximated to the second order by

$$(1) \quad W(h) \simeq W(m) + W'(m)(h-m) + 0.5\ W''(m)(h-m)^2$$

where m is the mean height of plants, and where W' and W'' are respectively the first and second derivatives of the height-weight function with respect to height. Taking the mathematical expectation of Equation (1), the following expression is obtained

$$(2) \quad E[W(h)] \simeq W(m) + 0.5\ W''(m)\ s^2$$

where s^2 is the variance of plant height.

This is a most interesting result. It says that expected weight may be calculated approximately by evaluating the weight function at the mean of height [as is done by NPI] and adding to this an amount that depends upon the value of the second derivative of the height-weight function evaluated at the mean height multiplied by the variance of plant height. Interpreted in another way, this result implies that the bias inherent in NPI's procedure is on the order of 0.5 $W''(m)\ s^2$.

The use of Equation (2) can be illustrated readily using the data assumed for the example shown in Figure 2. In particular,

$$W''(m) = (0.0754)^2(0.0095e^{0.0754(40)})=0.0011$$

and

$$s^2 = 0.5(60-40)^2+0.5(20-40)^2=400$$

so

$$\begin{aligned} E[W(h)] &\simeq 0.1939 + (0.5)(0.0011)(400) \\ &0.1939 + 0.2200 \\ &0.4139 \end{aligned}$$

This result [i.e., 0.4139] clearly is much nearer the true mean [recall from above that the true mean is 0.4594]. It is still a good ways off the mark, however, due to the facts that W(h) is more sharply curved than is reflected by a quadratic function and that the variance of plant height, s^2, is large. Note that the estimated bias in the NPI procedure [i.e. 0.2200] is larger than the value calculated by the procedure [i.e., 0.1939]. While this certainly is a result of the specific example chosen, it does illustrate the fact that the NPI procedure can result in extremely inaccurate estimates.

4 AN APPROXIMATE RECONCILIATION

The second order approximation developed in the preceding section along with data on the mean and variance of the height of the plants in the sample of plants NPI used to prepare Figure 1 can be used to explore possible reconciliation between the shrub density estimates presented in CONAZA (1977) and NPI (1980). The sample mean and variance of height of the plants in

NPI's sample are

$$m = 42.94 \text{ cm.}$$

$$s^2 = 219.34 \text{ cm.}^2$$

Using these estimates and the approximation given in Equation (2), it is calculated that

$$W(m) = 0.24 \text{ kg.}$$

$$0.5W''(m)s^2 = 0.15 \text{ kg.}$$

or that the average shrub weight in the sample was approximately 0.39 kilograms. An approximately correct estimate of average shrub weight is thus 1.62 times as large as the estimate that would be obtained by the NPI procedure [i.e. 0.39/0.24 = 1.62].

It is interesting to note that if 1.01 tons per hectare is taken as the average shrub density implied by the NPI study, the corresponding CONAZA (1977) estimate of 1.64 tons per hectare is, indeed, approximately 1.62 times the NPI estimate [i.e., 1.64/1.01 = 1.62]. Too much should not be made of this calculation. There are many other factors that could account for the differences [e.g., sampling error, basic differences in survey methods employed]. What is needed is a complete and systematic evaluation of methods and procedures employed in each study of guayule occurrence, and a critical effort to make the best possible estimate.

It can, however, be concluded that in subsequent field investigations in which rapid survey methods based on height-weight relationships [or, for that matter, other relationships] are employed, care should be taken in devising the estimating procedure to insure that it is not systematically biased. The procedure recommended by NPI is seriously biased although, as has been demonstrated above, it is a relatively simple matter to modify the procedure to eliminate this bias.

In any actual future application of the NPI technique [corrected as described above], a sample of guayule individuals should be taken and the height-weight relationship that is used to calculate plant weights should be re-estimated. Although the equation reported in Figure 1 appears to be a satisfactory characterization of the sample on which it is based, it is strongly recommended that an independent determination of the parameters of this relationship and its theoretical and empirical adequacy be made in each investigation in which the short-cut technique is to be employed.

REFERENCES

Anderson, R.J. 1981. An Estimate of the Maximum Sustainable Yield from Harvest of Guayule Wildstands in the Vicinity of Cedros. WP 81-57. International Institute for Applied Systems Analysis, Laxenburg, Austria. (May, 1981).

CONAZA. 1977. Estudio e Inventorio de las Poblaciones Naturales de Guayule (parthenium argentatum, gray) en la Region de Saltillo. (Coahuila: Comision Nacional de las Zonas Aridas).

Lloyd, F.E. 1911. Guayule (parthenium argentatum gray): A Rubber Plant of the Chihuahuan Desert (Washington, D.C.: The Carnegie Institution of Washington).

NPI. 1980. Ecology and Reproduction of Guayule (Parthenium argentatum, gray) in Natural Stands in Coahuila, Nuevo Leon, and Zacatecas, Mexico. A report prepared for the Centro de Investigacion en Quimica Aplicada.

NPI. 1981. Inventory of the Distribution and Standing Crop of Guayule (Parthenium argentatum, gray) in Northern Mexico. A draft report prepared for the Centro de Investigacion en Quimica Aplicada.

Rao, C.R. 1965. Linear Statistical Inference and Its Applications. (New York: John Wiley & Sons, Inc.).